*Indian J. Plant Physiol.*, Vol. 17, No. 2, (N.S.) pp. 121-127 (April-June, 2012)

# CLONING AND CHARACTERIZATION OF *MYO*- INOSITOL-1-PHOSPHATE SYNTHASE (*MIPS1*) GENE IN DEVELOPING SEEDS OF *GLYCINE MAX*

**AWADHESH KUMAR, SWETA KUMARI, MONICA JOLLY, R.D. RAI AND ARCHANA SACHDEV***

Division of Biochemistry, Indian Agricultural Research Institute, New Delhi-110012, India

Received on 1st December, 2011, Revised and accepted on 18th April, 2012

## SUMMARY

**Phytic acid (*myo*-inositol 1, 2, 3, 4, 5, 6 hexakisphosphate) is a phoshphorylated derivative of *myo*-inositol, which functions as a major storage form of phosphorous in plant seeds. *MIPS* (*myo*- inositol-1-phosphate synthase) catalyzes the first and rate limiting step in phytic acid biosynthesis. In the present study, a genomic sequence of *MIPS* from soybean (*Glycine max* var. Pusa16) was amplified using long PCR. Cloning and characterization of the genomic sequence revealed a total length of 2608 bp containing 9 introns interrupting 10 exons. The transcribed sequence of the gene had an expected open reading frame of 1533 bp encompassing *MIPS1a* variant of the *MIPS*1 isoform of *MIPS* gene. Strong homology with the previously reported *MIPS* gDNA as revealed in the BlastN search also indicated a high degree sequence similarity to *Phaseolus vulgaris, Vigna radiata*, *Cicer arietinum, Lotus japonica* and *Medicago truncatula*. The maximum level of phytate content observed in mature soybean seeds was 1.9%. *MIPS* expression by RT-PCR and Northern analysis revealed increased transcript levels during early stages of seed development reaching a maximum at 6-8 mm seed size. At least four copies of the *MIPS* were detected during Southern analysis of the genomic DNA.**

***Key words: GmMIPS1,* developing seeds, gene expression, phytate accumulation**

## INTRODUCTION

Soybean contains about 40% protein and 20% oil. It is predominantly grown as a sole *kharif* crop (June-October) in the northern, north-eastern, north-western and central parts of India. Almost 60-80% of total seed phosphorus in soybean is in the form of phytate (Raboy and Dickinson 1987), which is degraded by the enzyme activity of phytase during germination (Gibson and Ullah 1990). Soy meal, commonly used in animal feed because of its high protein content has high levels of phytic acid which have negative consequences in non-ruminant nutrition as they lack the digestive enzymes necessary for the efficient utilization of phytate phosphorous. Phytic acid acts as an antinutrient by chelating mineral cations and charged proteins, reducing their bioavailability and the overall digestibility of the feedstuff (Oatway *et al.* 2001). Nonruminants are unable to cleave phosphorus molecules from phytic acid, resulting in a requirement for expensive and non-renewable sources of inorganic phosphorus for optimal growth (Birch-Pedersen *et al.* 2002). Undigested phytate complexes pass through the gastrointestinal tract and are excreted in the manure. Nutrient rich manure when applied repeatedly to the croplands as fertilizer, results in environmental hazards as the soil microbes cleave phosphorous from phytic acid resulting in phosphorous loading in soil and phosphorus pollution in neighbouring watersheds leading to eutrophication of the aquatic ecosystem (Sharpley *et al.* 1994). Reducing the phytate levels in seeds can thus

*Corresponding author: arcs_bio@yahoo.com

increase the nutritional value of seeds and reduce the potential for phosphorus pollution. MIPS (*myo*-inositol-1-phosphate synthase), a highly conserved enzyme that has been identified throughout all biological kingdoms, catalyzes the NADH-dependent conversion of glucose-6-phosphate to *myo*-inositol-1-phosphate that represents the first step in phytic acid biosynthesis (Hegeman *et al.* 2001). This reaction is the only *de novo* source of inositol ring in all organisms. The conversion occurs early during seed development than accumulation of the final product, i.e. phytic acid. Because of its role in the initial synthesis of the phytic acid, the precursor, MIPS, becomes an attractive target for controlling phytate accumulation. Reducing the flux through the biosynthetic pathway by down regulating the activity of this key biosynthetic enzyme through gene silencing technology, may provide additional routes for future development of low phytic acid soybean phenotypes with improved phosphorus availability. The present investigation was carried out with the objective to clone and characterize the full length genomic sequence encoding *myo*-inositol-1-phosphate synthase from developing seeds of soybean (*Glycine max* L.) and also to study the changes in MIPS gene expression in different stages of seed development and other tissues types of soybean.

## MATERIALS AND METHODS

*Plant material and growth conditions:* Soybean seeds of *Glycine max* var. Pusa 16 were provided by Pulse Research Laboratory, Division of Genetics, Indian Agricultural Research Institute, New Delhi. The seedlings were harvested at 10 days after germination. Developing seeds of different sizes (2-4 mm, 4-6 mm, 6-8 mm, 8-10 mm, 10-12 mm and 12-14 mm) and other plant tissues types (flowers, roots, seedlings, leaves and stems), harvested from soybean grown under natural climatic conditions in the fields of IARI, were frozen immediately in liquid nitrogen and stored at -80$^0$C. The seeds were divided into groups based on the length as measured from apical end to basal end of the seed.

*Quantification of phytate***:** Samples were ground in liquid nitrogen in a pestle-mortar and the phytate content was determined and calculated by an assay procedure specific for the measurement of phosphorous from phytic acid, myo-inositol (phosphate) and monophosphate esters by phytase and alkaline phosphatase using Phytic Acid/ Total Phosphorus Assay Kit from Megazyme International Ireland Limited.

*Amplification of MIPS gene by long PCR-* Plant DNA from 6-7 days old seedlings was isolated by CTAB method as described by (Sambrook and Russell 1989). Long PCR-mediated amplification of the sequence flanked by the primers [FP:5'ATGTTCATCGAGAATT TAAGGTT3' and RP:5'ATCACTTGTACTCGAGAAT CAT3'] (Accession no. GQ471958) was performed using Phusion TM high-fidelity PCR Kit from Finnzymes. The mixture was treated at 98$^0$C (2 min) and subjected to 38 cycles of amplification, 98$^0$C (30 sec), 63.5$^0$C (30 sec), 72$^0$C (1 min 40 sec) with a final elongation of 10 min at 72$^0$C to generate the amplicon.

*Cloning, sequencing and molecular characterization:* The PCR amplified genomic DNA sequence was cloned in pBluescript SK$^+$ vector (3.0 kb) and sequenced by an automated sequencer. Standard techniques were used for DNA cloning and Southern and Northern analysis (Sambrook and Russell 1989).

*RT-PCR:* 5 µg of total RNA was used to prepare total cDNA by two step RT-PCR using the RevertAid™ H Minus First Strand cDNA Synthesis Kit (MBI Fermentas). RT–PCR was carried out with *GmMIPSF* [5'-ACCAGGGCTGATTGATCTTG-3'] and *GmMIPS R* [5'TTGGAGATTTCCTTGGAACG-3'] (Accession No. GQ477134) in a 25 µl aliquot containing 20 ng of cDNA, 2.5 µl of 10X Taq buffer, 2 mM $MgCl_2$, 1.0 µl of 10mM dNTP mix, 200 nM of each primer and 0.5 U of Taq DNA polymerase (MBI Fermentas). The mixture was treated at 94$^0$C for 2 min and subjected to 36 cycles of amplification (94$^0$C for 30 sec, 57.6$^0$C for 30 sec, 72$^0$C for 30 sec) with a final elongation cycle of 10 min at 72$^0$C. The 213 bp amplicon was cloned in the pGEM®-T Easy vector (Promega).

*Southern analysis:* DNA samples (10 µg) were restricted and resolved on an agarose gel and transferred onto a nylon membrane (Amersham). For use as a hybridization probe, the *MIPS* specific amplicon (213 bp) was gel purified and labeled with α-[$^{32}$P]dATP (3000

Ci/mol) by random oligo priming (HexaLabel™ DNA Labeling Kit, MBI Fermentas). High stringency DNA hybridization and subsequent washes were performed according to (Sambrook and Russell 1989). The blot was further dried and exposed to an X-Ray film (Kodak).

*Northern analysis:* Total RNA was extracted using the RNAeasy Kit (Qiagen, Valencia, CA, USA) from 100 mg of developing seeds and from other tissue types (young roots, stems, leaves, flowers and seedlings). Quality and quantity of total RNA were estimated by gel visualization and Specord-200 UV/Vis Spectrophotometer (Analytic Jena AG.). Northern blotting was carried out with 10 µg of RNA as described by Sambrook and Russell 1989. The 213 bp RT-PCR amplicon was gel purified and labeled for use as a hybridization probe.

## RESULTS AND DISCUSSION

*Phytate accumulation*: Phytic acid accumulation was linear throughout seed development. It was detected early during embryogenesis in field grown soybean and accumulated in a linear manner in the developing seeds until physiological maturity. Starting from 0.29 g/100 g DW in 2-4 mm seeds the phytate level reached its maximum to 1.9 g/100 g DW in the mature seeds (12-14 mm) (Fig. 1). The increase in the accumulation was however less pronounced during early stages of development. Raboy and Dickinson (1987) measured phytate and phosphate levels in the developing seeds and showed that phytic acid content increased steadily until late in seed maturation (18.7 to 33.6 µg per seed per day). Low phytate concentration in the present study at the initial stages of seed development probably coincides with the period of intense seed metabolism.

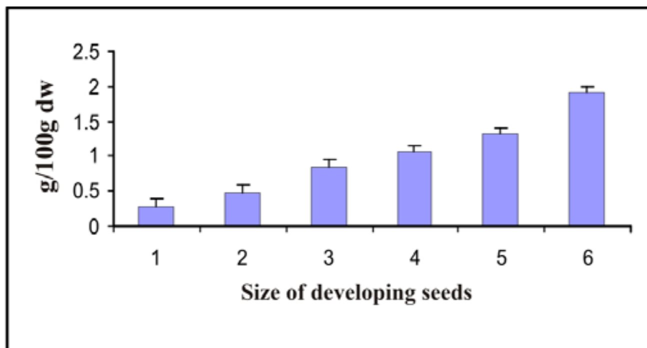



**Fig. 1. Changes in phytate concentration in developing seeds of soybean on dry weight basis. The data represents the mean (± S.D) for triplicate samples.**

Estimation of phytate content in other tissue types of soybean (leaves, flowers, roots, stems and seedlings (Fig. 2). presented lower amounts compared to mature cotyledons indicating that phytic acid is the primary storage compound in the seeds (70-80%) and that phytate deposition is mainly restricted to cells that remain active during the quiescent phase of seed development.

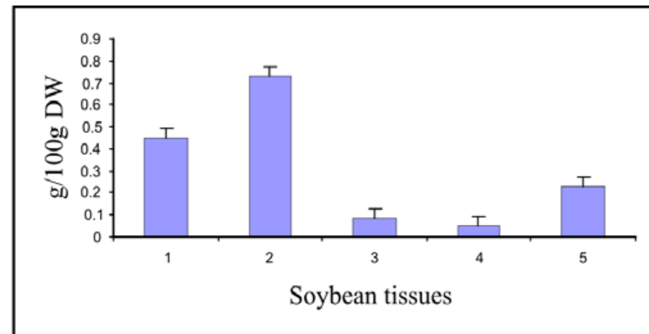



**Fig. 2. Changes in phytate concentration in soybean tissues on dry weight basis. The data represents the mean (± S.D) for triplicate samples.**

*Molecular characterization of MIPS*: Using the sequence data of the soybean *MIPS* genomic DNA, available in NCBI GenBank (GenBank Acc. No. DQ323904; Chappell *et al.* 2006), a set of primers was designed for the 5' and 3' end regions and a 2.608 kb gDNA was amplified using the long PCR protocol. The amplicon was cloned in *p*Bluescript $SK^+$ vector by blunt-end ligation and sequenced (Fig. 3). The nucleotide sequences were compared to the non-redundant databases using a BlastN search. A 100% homology at 100% coverage with previously reported *Glycine max MIPS* gDNA (Acc. No. DQ323904) and 100% homology at 60% coverage with *Glycine max MIPS* cDNA (Acc. No. AF293970) indicated the presence of introns in between the coding regions. The alignment of *MIPS* gDNA sequence with the *MIPS* cDNA fragment in BLASTN results showed the presence of 10 exons and 9 introns and their specific locations (Fig. 4). The nucleotide composition of the sequence was deduced using the BioEdit software. Molecular weight of 1580.986 KDa and a G+C content of 39.95% and A+T content of 60.05% was deduced. An open reading frame

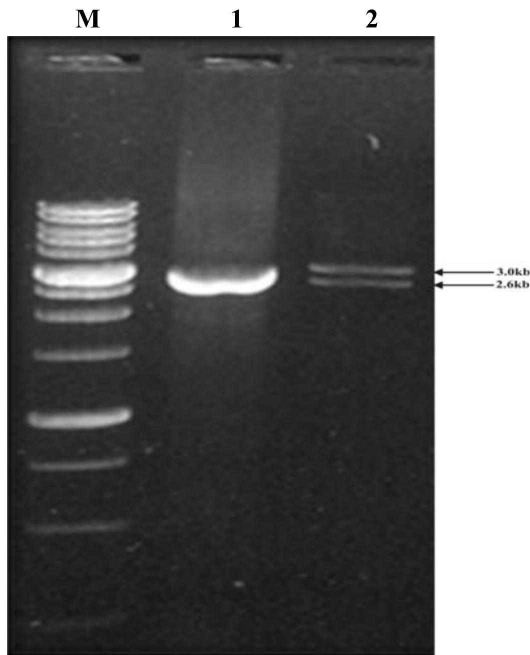
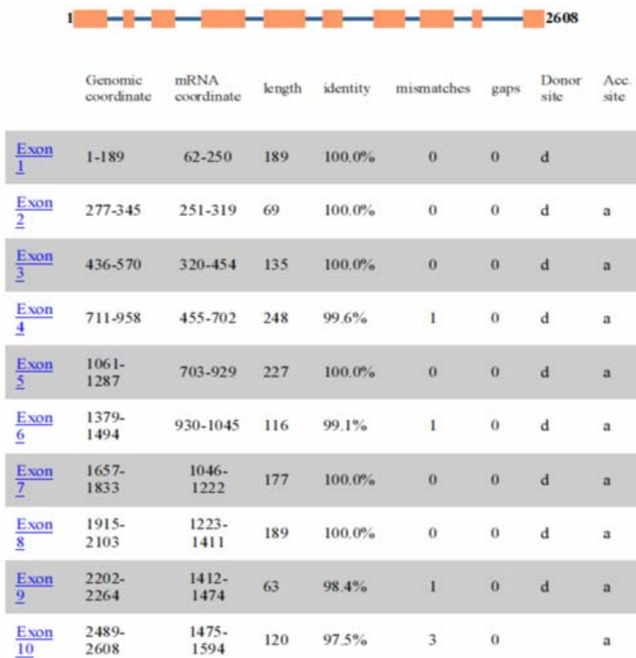



**Fig. 3. 1.0% Agarose gel showing long PCR amplified *MIPS* fragment (Lane-1), EcoRI+HindIII λ DNA marker (Lane M), rrestriction analysis of *MIPS* amplicon (~2.5kb) cloned in SK+ vector (3kb) (Lane 2).**




| | Genomic coordinate | mRNA coordinate | length | identity | mismatches | gaps | Donor site | Acc. site |
|---|---|---|---|---|---|---|---|---|
| Exon 1 | 1-189 | 62-250 | 189 | 100.0% | 0 | 0 | d | |
| Exon 2 | 277-345 | 251-319 | 69 | 100.0% | 0 | 0 | d | a |
| Exon 3 | 436-570 | 320-454 | 135 | 100.0% | 0 | 0 | d | a |
| Exon 4 | 711-958 | 455-702 | 248 | 99.6% | 1 | 0 | d | a |
| Exon 5 | 1061-1287 | 703-929 | 227 | 100.0% | 0 | 0 | d | a |
| Exon 6 | 1379-1494 | 930-1045 | 116 | 99.1% | 1 | 0 | d | a |
| Exon 7 | 1657-1833 | 1046-1222 | 177 | 100.0% | 0 | 0 | d | a |
| Exon 8 | 1915-2103 | 1223-1411 | 189 | 100.0% | 0 | 0 | d | a |
| Exon 9 | 2202-2264 | 1412-1474 | 63 | 98.4% | 1 | 0 | d | a |
| Exon 10 | 2489-2608 | 1475-1594 | 120 | 97.5% | 3 | 0 | | a |

**Fig. 4. Alignment as on plus strand of *MIPS* genomic sequence and on plus strand of mRNA sequence (Accession No. AF293970.1), mRNA coverage: 86% Overall percent identity: 99.6%, Non-aligning poly(A) tail: 23.**

of 1533 bp in length was predicted which was found to be similar in length to the previously reported *Gm MIPS 1a* gene sequence by Hegeman *et al.* 2001 and Chappell *et al.* 2006 indicating that amplified *MIPS* fragment encoded the complete ORF region of *MIPS*. Alignment on the plus strand of *MIPS* genomic sequence (Accession No. DQ323904) and on the plus strand of mRNA sequence (Acc. no. AF293970) showed 99.6% identity at 86% coverage respectively. The genomic sequence was predicted to encode a 510 amino acid protein using the FGENESH program of *Soft berry* web server. BLAST database searches were performed with the *MIPS* cDNA sequence and the predicted protein sequence (Altschul *et al.* 1997). The search revealed a high degree of sequence identity between *Gm MIP1a* and *MIPS* genes from plants and other organisms. It showed high score identities with *Phaseolus vulgaris* (95%), *Cicer arietinum* (94%), *Nicotiana tabacum* (92%), *Arabidopsis thaliana* (87%), *Brassica napus* (88%), *Zea mays* (88%), *Oryza sativa* (88%) and *Tritcum aestivum* (88%). Alignment of the conserved domains of MIPS protein in various plants revealed seven conserved regions from 59-78 position (Domain I), 118-134 position (Domain II), 301-367 position (Domain III), 376-390 position (Domain IV), 394-411 position (Domain V), 417-439 position (Domain VI) and 494-509 position (Domain VII) using the BioEdit software (Fig. 5). Genomic DNA was digested with *BamH*1, *EcoR*1, *EcoR*V and *Hind*III separately and probed with the 213bp internal fragment represented in the Domain III of *MIPS* cDNA. Southern blot analysis under high stringency revealed multiple bands of high molecule weight ranging in size from 3kb to 23 kb (Fig. 6). The data confirmed the presence of four loci with *MIPS* homology as previously reported by Hegeman *et al.* 2001. Restriction fragments generated in the *EcoR*1 and *Hind*III lanes were larger than the PCR amplified product confirming the lack of these sites in the *MIPS* genomic sequence. In maize, seven homologous *MIPS* sequences were mapped on different chromosomes (Larson and Raboy 1999), one gene in *Passiflora edulis* f. *flavicarpa* (Abreu and Aragao 2007) and two distinct *MIPS* genes were identified in Arabidopsis (Johnson and Burk 1995, Johnson and Sussex 1995). Multiple ESTs with homology to *MIPS* sequences from various soybean tissues types have been identified in databases suggesting

**Fig. 5. Alignment of deduced amino acid sequences of various plant *MIPS* showing seven conserved domains (shown in boxes).**

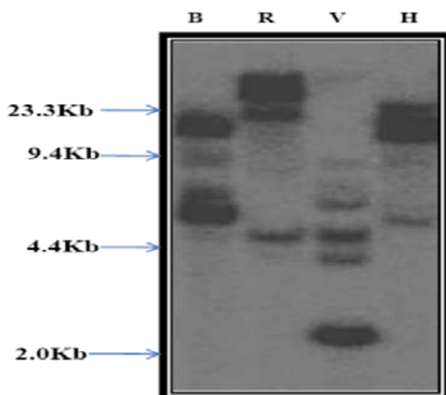



**Fig. 6. Southern-blot analysis of soybean genomic DNA. DNA samples (10 µg) digested with *Bam*HI (B), *Eco*RI (R), *Eco*RV (V) and *Hin*dIII (H) respectively and probed with $^{32}$P labelled *MIPS* *c*DNA under high stringency.**

that multiple bands observed under stringent conditions during Southern analysis could correspond to these highly similar sequences and that *MIPS* sequence amplified in the present study, *GmMIPS1*, represents one member of the multigene family.

*Expression analysis*: Northern and RT-PCR analysis with the *MIPS* specific probe revealed the presence of *MIPS* transcripts of ~1800 nucleotides in length in the early stages of seed development i.e. 2-4 mm seed size. As the development progressed the transcripts attained maximal expression level up to 6-8 mm of seed size stage followed by a gradual decline to nearly undetectable levels in 12-14 mm seeds (Fig. 7) and (Fig. 9). Steady state RNA levels were higher in developing seeds than in other tissue types including flowers, stems, roots and seedlings (Fig. 8) where the complete absence of *MIPS* transcripts was significant thereby suggesting a peak demand for *myo*-inositol during mid stages (6-8 mm) of seed development for synthesis of various metabolites for which *myo*- inositol acts as precursor.

The spatial and temporal expression profiles observed in the present study could thus be well correlated with the earlier findings of Hegeman *et al.* 2001 as they indicated that high steady state levels of *MIPS* expression should precede and/or accompany the synthesis and accumulation of phytic acid since phytate levels were observed to increase at a linear rate until

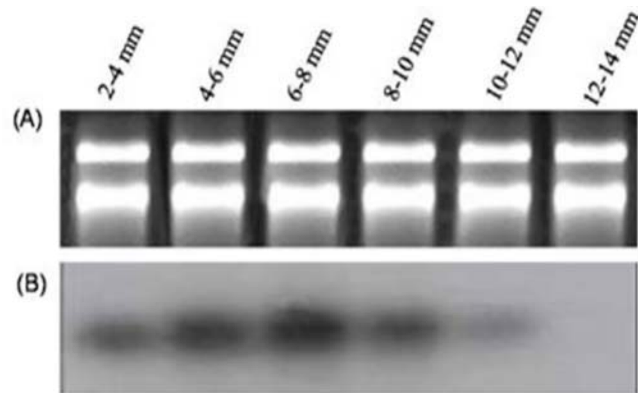



**Fig. 7. Northern blot analysis of *MIPS* transcripts from developing seeds of soybean. Samples of total RNA after transfer to membrane were probed with $^{32}$P labeled *MIPS* cDNA. (A) Total RNA from 2-4 mm, 4-6 mm, 6-8 mm, 8-10 mm, 10-12 mm and 12-14 mm seed size as loading control (B) Autoradiograph of the samples.**

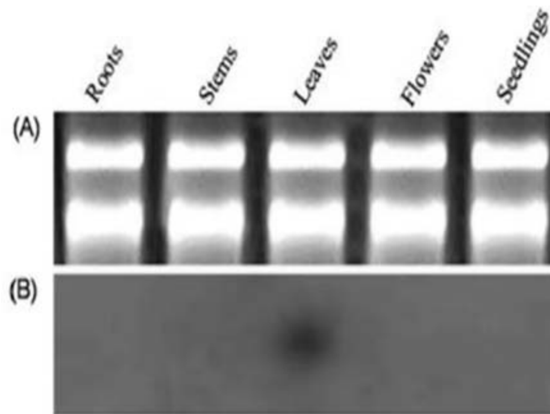



**Fig. 8. Northern blot analysis of *MIPS* transcripts from vegetative tissues of soybean. Samples of total RNA transfered to membrane were probed using $^{32}$P labeled *MIPS* cDNA. (A) Total RNA from roots, stems, leaves, flowers and seedlings as loading control (B) Autoradiograph of the samples.**

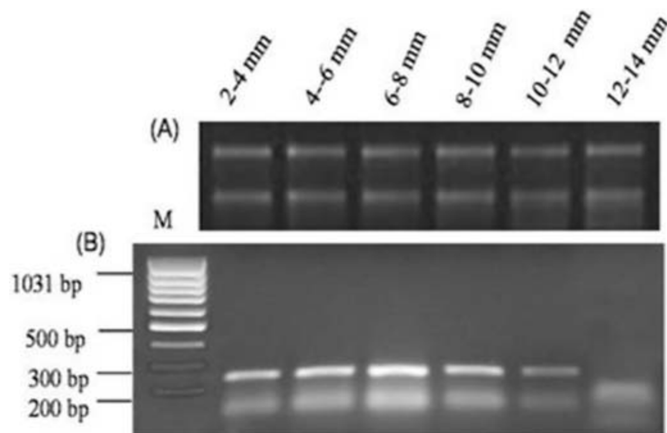



**Fig. 9. Differential transcription of the *GmMIPS* gene from developing seeds through RT- PCR analysis. (A) Total RNA from 2-4 mm, 4-6 mm, 6-8 mm, 8-10 mm, 10-12 mm, 12-14 mm developing seeds sizes to assess the quality and quantity. (B) RT-PCR expression analysis, in the presence of 213 bp *MIPS* cDNA specific primer pair, M - 100 bp DNA ladder.**

late in seed maturation. Negligible or low basal levels of transcripts in other tissues besides seeds suggested a relatively low demand for inositols in them and also the low expression levels in these tissues are probably sufficient for the production of inositol and inositol phosphates involved in the other aspects of inositol metabolism required for the growth and development. *In silico* northern analysis generated from soybean EST database predicted an expression of four highly similar *MIPS* coding sequence where ESTs from *Gm MIP3* were observed only in leaf and bud libraries and ESTs from *Gm MIPS1* were observed only in immature cotyledons (Hegeman *et al.* 2001). Screening data from 15 EST libraries by Hitz *et al.* 2002 also showed *Gm MIPS1* to be the preferred gene of the seed although expressed at lower levels in other tissue types, while *GmMIPS2* was expressed in many tissues but not in developing seeds.

## CONCLUSION

The RNA expression data in the present study thus supports the organ specific expression of *MIPS1* gene controlled by a strong and developmentally regulated promoter in immature seeds and its abundant expression in developing seeds suggests its critical role in PA biosynthesis.

*Accession number*: Sequenced data from these studies have been deposited in GenBank with Accession no. GQ471958 for 2.608 kb genomic DNA sequence, GQ477134 for 213 bp cDNA sequence.

## ACKNOWLEDGEMENT

Financial support provided by ICAR, Govt. of India, in the form of JRF is duly acknowledged.

## REFERENCES

Abreu, E.F.M. and Aragão, F.J.L. (2007). Isolation and characterization of a *myo*-inositol-1-phosphate synthase Gene from yellow passion fruit (*Passiflora edulis* f. *flavicarpa*) expressed during seed development and environmental stress. *Ann Bot.* **99:** 285–292.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucl. Acids Res.* **25:** 3389-3402.

Birch-Pedersen, H., Sorensen, L.D. and Holm, P.B. (2002). Engineering crop plants: getting a handle on phosphate, *Trends Plant Sci.* **7:** 118-125.

Chappell, A.S., Scaboo, A.M., Wu, X., Nguyen, H., Pantalone, V.R. and Bilyeu, K.D. (2006). Characterization of the

*MIPS* Gene family in *Glycine max. Plant Breed*. **125:** 493-500.

Gibson, D.M. and Ullah, A.H.J. (1990). Phytases and their action on phytic acid. In: D.J. Morre, W.F. Boss, F.A. Loewus, Inositol Metabolism in Plants, *Wiley-Liss, Inc*., New York. pp. 77-92.

Hegeman, C.E., Good, L.L. and Grabau, E.A. (2001). Expression of D-myo-inositol-3-phosphate synthase in soybean, Implications for phytic acid biosynthesis, *Plant Physiol.* **125:** 1941-1948.

Ishitani, M., Majumder, A.L., Bornhouser, A., Michalowski, C.B., Jensen, R.G. and Bohnert, H.J. (1996). Coordinate transcriptional induction of *myo*-inositol metabolism during environmental stress. *Plant J.* **9:** 537–548.

Johnson, M.D. and Burk, D. (1995). Isozyme of 1L-*myo*-inositol-phosphate synthase from Arabidopsis (Accession No. U30250); PGR95-067). *Plant Physiol*. **109:** 721.

Johnson, M.D. and Sussex, I.M. (1995). 1-L-*myo*-inositol 1-phosphate synthase from *A. thaliana. Plant Physiol.* **107:** 613- 619.

Larson, S.R. and Raboy, V. (1999). Linkage mapping of maize and barley *myo*-inositol 1-phosphate synthase DNA sequences: correspondence with a *low phytic acid* mutation. *Theor. Appl. Genet.* **99:** 27-36.

Majumder, A.L., Johnson, M.D. and Henry, S.A. (1997). Structural studies of MIP synthase *Biochim. Biophys. Acta.* **1348:** 245–256.

Oatway, L., Vasathan, T. and Helm, J.H. (2001). Phytic acid. *Food Reviews International.* **17:** 419-431.

Raboy, V. and Dickinson, D.B. (1987). The timing and rate of phytic acid accumulation in developing soybean seeds. *Plant Physiol.* **85:** 841-844.

Sambrook, J. and Russell, D.W. (1989). *Molecular cloning: A laboratory manual,* Cold Spring Harbor Laboratory press.

Sharpley, A.N., Chapra, S.C., Wedepohl, R., Sims, J.T., Daniel, T.C. and Reddy, K.R. (1994). Managing agricultural phosphorus for protection of surface waters: Issues and Options. *J. Environ. Qua.* **23:** 437-451.